



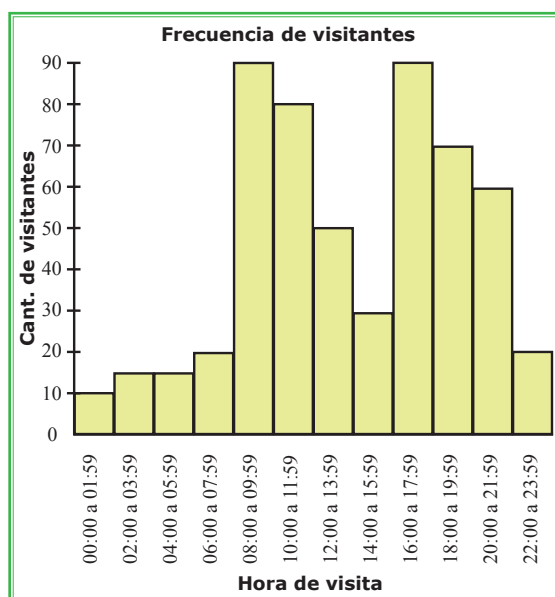
Capítulo 11

Estadística y Probabilidades

Introducción

Generalmente la estadística se asocia a un conjunto de datos organizados en tablas o en gráficos, referentes a geografía, demografía, economía, mercados, salud, entre otros temas. Pero la estadística es mucho más amplia de lo que parece. Es una ciencia tan antigua como la matemática, y por su utilidad, es apoyo de todas las demás ciencias. La estadística ayuda a que los administradores tomen las mejores decisiones en tiempos de incertidumbre.

La historia de la estadística se resume en tres grandes etapas, a saber:



Primera Etapa (Los Censos): La necesidad de llevar un control de la población y las riquezas existentes en un reino o país es una de las primeras aplicaciones de la estadística. De acuerdo al historiador griego Herodoto, los faraones lograron recopilar hacia el año 3050 a.C. datos relativos a la población y la riqueza de Egipto. Esta etapa se caracterizó por ser netamente descriptiva de la información.

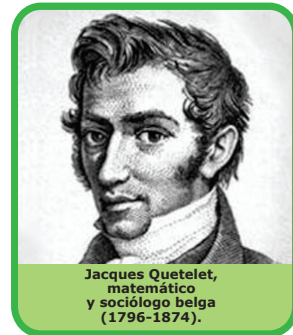
Segunda Etapa (De la descripción de los conjuntos a la aritmética política): Luego de superar la etapa descriptiva, surgieron obras como las de Graunt, Petty y Halley. El penúltimo es autor de la famosa obra *“Aritmética Política”*. Chaptal, ministro francés del interior, publicó en 1801 el primer censo general de población y desarrolló estudios industriales sobre producción, haciéndose sistemáticos durante las dos terceras partes del siglo XIX.

Tercera Etapa (Cálculo de Probabilidades): El cálculo de probabilidades surge como una poderosa herramienta de análisis para estudiar fenómenos económicos y sociales, y otros cuyas causas son difíciles de controlar.

Durante cierto tiempo, la teoría de las probabilidades limitó su aplicación a los juegos de azar, y hasta el siglo XVIII no comenzó a aplicarse a los grandes problemas científicos.

La estadística que conocemos hoy en día, debe gran parte de su realización a los trabajos matemáticos de aquellos hombres que desarrollaron la teoría de las probabilidades, con la cual se adhirió la estadística a las ciencias formales.

De aquí, que la estadística es la ciencia que recolecta, organiza, describe e infiere información para el análisis e interpretación de los fenómenos que nos rodean.



11.1 Estadística descriptiva

Objetivos

Al finalizar esta sección el lector podrá:

- * Explicar el rol de la estadística en la sociedad y su aplicación en el análisis de información.
- * Distinguir entre estadística descriptiva y estadística inferencial.
- * Identificar los errores más comunes cuando se analiza información estadística.
- * Definir los términos estadísticos, los tipos de variables y escalas de medición frecuentemente más empleados.

La estadística es la ciencia cuyo objetivo es reunir información cuantitativa concerniente a individuos, grupos, series de hechos, etc., para deducir de su análisis, conclusiones precisas o previsiones para el futuro.

Algunos confunden comúnmente los términos asociados con la estadística, debido a que esta palabra etimológicamente tiene tres significados: en primer término se usa para referirse a la información estadística; también se utiliza para referirse al conjunto de técnicas y métodos que se utilizan para analizar la información estadística; y el término estadístico, en singular y en masculino, se refiere a una medida derivada de una muestra.

Estadística y Probabilidades

Para su mejor estudio, la estadística se ha dividido en dos grandes ramas: la estadística descriptiva y la estadística inferencial.

La **estadística descriptiva** consiste en la presentación de datos en forma de tablas y gráficas. Ésta comprende cualquier actividad relacionada con los datos y está diseñada para resumir o describir los mismos sin factores pertinentes adicionales; esto es, sin intentar inferir nada que vaya más allá de los datos como tales.

Los métodos estadísticos tradicionalmente se utilizan para propósitos descriptivos, para organizar y resumir datos numéricos. La estadística descriptiva trata de la tabulación de datos, su presentación en forma gráfica o ilustrativa y el cálculo de medidas descriptivas.

Estas técnicas estadísticas se aplican de manera amplia en control de calidad, contabilidad, mercadotecnia, estudios de mercado, análisis deportivos, administración de instituciones, educación, política, medicina, y por aquellas personas que intervienen en la toma de decisiones.

La **estadística inferencial** se deriva de muestras, que son subconjuntos de una población con alguna característica de interés. A partir de las observaciones hechas a una parte de un conjunto numeroso de elementos, se infiere acerca de las características que posee la población. Esto implica que su análisis requiere de generalizaciones que van más allá de los datos.

Como consecuencia, la característica más importante del reciente crecimiento de la estadística ha sido un cambio en el énfasis de los métodos que sirven para hacer generalizaciones. La estadística inferencial investiga o analiza una población partiendo de una muestra tomada.

El **método estadístico** es el conjunto de los procedimientos que se utilizan para medir las características de los datos, para resumir los valores individuales y para analizarlos, a fin de extraerles el máximo de información; es lo que se conoce como método estadístico.

Un método estadístico contempla las siguientes seis etapas:

1. Definición del problema.
2. Recopilación de la información existente.
3. Clasificación y control de calidad de los datos.
4. Codificación y digitación.
5. Análisis.
6. Presentación.

Errores estadísticos comunes. Existe la posibilidad de cometer errores al momento de recopilar los datos que serán procesados, así como durante el cómputo de los mismos. No obstante, hay otros errores que no tienen que ver con la digitación y no son tan fáciles de identificar.

Algunos de estos errores son:

- **Sesgo:** Hay que ser completamente objetivo y no tener ideas preconcebidas antes de comenzar a estudiar un problema, evitando que puedan influir en la recopilación y en el análisis de la información. En estos casos se dice que hay un sesgo cuando el individuo da mayor peso a los datos que apoyan su opinión, que a aquellos que la contradicen. Un caso extremo de sesgo sería la situación donde primero se toma una decisión y después se utiliza el análisis estadístico para justificar la decisión ya tomada.
- **Datos no comparables:** Establecer comparaciones es una de las partes más importantes del análisis estadístico, pero es extremadamente importante que tales comparaciones se hagan entre datos que se presten a ello.
- **Proyección descuidada de tendencias:** La proyección simplista de tendencias pasadas hacia el futuro, es uno de los errores que más ha desacreditado el uso del análisis estadístico. Hay que tener en cuenta que cualquier estadística que se realice es una fotografía instantánea, por lo cual deben emplearse los históricos de otras estadísticas previamente realizadas antes de emitir cualquier resultado concluyente sobre el fenómeno o hecho que se está estudiando.
- **Muestreo incorrecto:** En la mayoría de los estudios, la información disponible es tan extensa que se hace necesario inferir a partir de muestras, para derivar conclusiones acerca de la población a la que pertenece la muestra. Si la muestra se selecciona correctamente, tendrá básicamente las mismas propiedades que la población de la cual fue extraída; pero si el muestreo se realiza incorrectamente, entonces puede suceder que los resultados no sean representativos de la realidad poblacional.

Cabe anotar que en este texto nos limitaremos a tratar la estadística descriptiva, para lo cual definiremos algunos conceptos básicos asociados a esta rama.

Conceptos básicos

Elemento o ente: Cualquier elemento que aporte información sobre la característica que se estudia. Así, si estudiamos la altura de los niños de una clase, cada alumno es un ente; si estudiamos el precio de la vivienda, cada vivienda es un ente.

Estadística y Probabilidades

Población: Conjunto o colección de los entes de interés. Cada ente presenta características determinadas, observables y medibles. Por ejemplo, en el elemento persona: nombre, edad, género, peso, nacionalidad, etc. Por lo tanto, la estadística se preocupa de estudiar las características de los elementos constituyentes de la población, y estudia las posibles relaciones y las regularidades que presenta la población a partir de estas características.

La población se puede clasificar, según su tamaño, en dos tipos:

- **Población finita:** El número de elementos es finito. Por ejemplo: la cantidad de alumnos de una escuela.
- **Población infinita:** El número de elementos es infinito o tan grande que pueden considerarse en cantidad infinita. Por ejemplo: las estrellas de la Vía Láctea.

Muestra: La mayoría de los estudios estadísticos, no se realizan sobre la población por los altos costos en tiempo y dinero, sino sobre un subconjunto o una parte de ella denominada **muestra**, partiendo del supuesto de que este subconjunto presenta el mismo comportamiento y características de la población. Por ejemplo, para la población “estudiantes de las escuelas de Guayaquil”, una muestra podría ser “el conjunto de niños de una escuela en particular”.

Variable: Es una característica que se asocia a los elementos de una muestra o población. Tiene la propiedad de poder ser medida u observada. Su expresión numérica es el **dato**. Las variables se pueden clasificar en dos tipos:

Variables cuantitativas: Se expresan por medio de números y pueden ser:

- **Discretas:** Sólo se miden por medio de valores puntuales. Por ejemplo: número de materias, cantidad de médicos en un hospital; y,
- **Continuas:** Pueden tomar cualquier valor intermedio entre dos números, es decir, intervalos. Por ejemplo: el peso y la estatura de una persona.

Variables cualitativas o atributos: No se pueden expresar numéricamente, sino por medio del nombre de la característica en estudio; se pueden clasificar en:

- **Ordinales:** Aquellas que sugieren una ordenación. Por ejemplo: nivel de estudio, posición de los ganadores de un concurso; y,

- **Nominales:** Aquellas que sólo admiten una mera ordenación alfabética, pero no establecen orden por su contenido. Por ejemplo: género, estado civil, color de cabello.

Las variables también se pueden clasificar en:

- **Variables unidimensionales:** Sólo recogen información sobre una característica. Por ejemplo: edad de los alumnos de una clase.
- **Variables bidimensionales:** Recogen información sobre dos características de la población. Por ejemplo: edad y estatura de los alumnos de una clase.
- **Variables multidimensionales:** Recogen información sobre tres o más características. Por ejemplo: edad, estatura y peso de los alumnos de una clase.

Existen otros tipos de variables que se emplean en estadística, como son las variables independientes, variables dependientes, variables explicativas, variables de respuesta, entre otras, que no forman parte del estudio de este capítulo, pero que se pueden analizar con más detenimiento en un curso formal de estadística.

Antes de realizar un estudio estadístico, es importante tener claridad respecto de aspectos tales como: ¿qué se está midiendo?, ¿cómo se está midiendo?, ¿para qué se está midiendo?, ¿por qué se está midiendo?, ¿quién está midiendo?, ¿en qué momento se realizarán las mediciones? En síntesis, el proceso de medición genera el tipo de variable y no su característica o propiedad por la cual está siendo estudiada.

Escala de medición de variables

Entenderemos por **medir**, el proceso que se realiza al asignar números a objetos, fenómenos o características, según ciertas reglas de cuantificación. Se asocia un número y sólo uno (correspondencia biunívoca) al objeto o fenómeno que se quiere medir.

Toda **medida** de una variable está referida a una escala, que es una regla o patrón que permite asociar números a los objetos o fenómenos. Los tipos de medida están representados en el siguiente cuadro:

ESCALA	DEFINICIÓN OPERACIONAL	EJEMPLO
Nominal	Las observaciones del atributo de la variable son clasificadas en categorías, solamente se puede verificar la igualdad entre las categorías. Nombran, pero no miden la variable.	Género: masculino/femenino. Estado civil: casado/soltero/divorciado/ viudo/unión libre.

Estadística y Probabilidades

Ordinal	Se pueden establecer relaciones de orden entre los datos de la variable: mayor, menor o igual. Contiene a la escala nominal.	Nivel socio económico: alto/medio/bajo. Rendimiento académico: excelente/regular/deficiente.
De intervalo	Ordenan las medidas y permiten realizar comparaciones entre dos medidas. Usan un cero relativo.	Temperatura: 10 °C, 12 °F. El rendimiento en una prueba puede estar en el intervalo [0, 20].
De razón	Se pueden establecer razones entre los datos. Posee cero absoluto.	Peso, estatura.

11.2 Organización de los datos

Objetivos

Al finalizar esta sección el lector podrá:

- * Dado un conjunto de datos, organizar la información empleando tablas de frecuencia.
- * Interpretar información estadística en forma tabular a nivel de frecuencia relativa y frecuencia acumulada.

Una vez que las variables han sido medidas, estos valores pasan a constituir el conjunto de datos estadísticos, que deberán ser procesados y analizados por el investigador. Para el efecto, existen algunas técnicas que serán estudiadas en esta sección.

Tablas de frecuencias

Según el número de observaciones y el rango de la variable, podemos clasificar las tablas de la siguiente manera:

Tablas de tipo I: El tamaño de la población o muestra es pequeño. Por ejemplo, las edades de 6 personas: 15, 18, 19, 21, 24, 28. Sólo se ordenan de manera creciente o decreciente.

Tablas de tipo II: El tamaño de la población o muestra es grande y el rango de la variable es pequeño.

Ejemplo 11.1 Tabla de tipo II.

El número diario de llamadas telefónicas realizadas en una casa durante 30 días, se encuentra tabulado así:

2	4	1	3	2	5
3	1	3	4	1	1
1	5	3	1	2	3
2	1	5	3	4	2
3	4	1	2	5	5

Sea la variable el número diario de llamadas telefónicas, podemos observar que el rango de la variable está entre 1 y 5 llamadas, y que el total de datos es 30 llamadas. Por lo tanto, la tabla de frecuencia se estructura siguiendo los pasos 1 y 2:

1. Ordene los datos en forma decreciente o creciente por cada columna y realice el conteo:

1	1	2	3	3	5
1	1	2	3	4	5
1	1	2	3	4	5
1	2	2	3	4	5
1	2	3	3	4	5



Nº. de llamadas	Conteo
1	
2	
3	
4	
5	

2. Estructure la tabla de frecuencia relacionando el conteo con un número (frecuencia):

Nº. de llamadas	Frecuencia
1	8
2	6
3	7
4	4
5	5
Total	30

Estadística y Probabilidades

Tablas de tipo III (Tabla de intervalos): El tamaño de la población o muestra es grande y el rango de la variable es grande.

Ejemplo 11.2 Tabla de tipo III.

La edad de un grupo de 30 personas se encuentra tabulada así:

22	23	44	10	28	40
15	43	38	7	24	31
28	12	5	20	18	47
50	27	14	16	30	26
55	27	42	50	27	36

Sea la variable, la edad de las personas, observamos que los valores están dispersos y que el rango de la variable está entre 5 y 55, por lo cual, si se quiere elaborar una tabla, ésta debe ser de intervalos. Ahora, los pasos que se deben seguir son:

1. Determine el total de datos. En este caso $N = 30$.
2. Calcule el rango R de la variable con la expresión $R = X_{\max} - X_{\min}$, en los cuales están considerados el valor máximo y mínimo de dicha variable. Para el ejemplo, $R = 55 - 5 = 50$.
3. Determine el número de intervalos, entre 10 y 15. En este ejemplo, se tomarán 13 intervalos.
4. Calcule la amplitud de los intervalos $i = \frac{R}{\text{Nº. Intervalos}}$, aproximando al entero más cercano. Para el ejemplo, $i = \frac{50}{13} \approx 4$.
5. Construya la tabla considerando que los intervalos serán siempre cerrados por la izquierda y abiertos por la derecha $[L_{i-1}, L_i)$. Para el primer intervalo $[L_1, L_2)$, L_1 es el mínimo valor de los datos y L_2 es igual a $L_1 + i$. Para el segundo intervalo $[L_2, L_3)$, L_2 ya se determinó en el paso anterior y L_3 es igual a $L_2 + i$. Este procedimiento se sigue realizando para los nuevos intervalos.

Para el ejemplo, la tabla sería:

Intervalos de edades	Frecuencia
[5, 9)	2
[9, 13)	2
[13, 17)	3
[17, 21)	2
[21, 25)	3
[25, 29)	6
[29, 33)	2
[33, 37)	1
[37, 41)	2
[41, 45)	3
[45, 49)	1
[49, 53)	2
[53, 57]	1
Total	30

Tablas de distribución de frecuencias: Generalmente, las tablas de tipo II y III se completan con distintos tipos de frecuencias, tales como:

- a) **Frecuencia absoluta:** Es el número de veces que aparece dicho valor, como resultado de la medición de la variable. Se denota por f_i .
- b) **Frecuencia absoluta acumulada:** Es el resultado de sumar a la frecuencia absoluta del valor correspondiente la frecuencia absoluta del valor anterior. Se denota por F_i .
- c) **Frecuencia relativa:** Es el cociente entre la frecuencia absoluta y el tamaño de la muestra o población: $h_i = \frac{f_i}{N}$, donde N = Tamaño de la muestra o población. Se denota por h_i .
- d) **Frecuencia relativa acumulada:** Es el resultado de sumar a la frecuencia relativa del valor correspondiente la frecuencia relativa del valor anterior. Se denota por H_i .

Estadística y Probabilidades

Modelos de tablas estadísticas

Tablas de tipo II (Variable cuantitativa discreta)

Variable	Frecuencia absoluta (f_i)	Frecuencia absoluta acumulada (F_i)	Frecuencia relativa (h_i)	Frecuencia relativa acumulada (H_i)
X_1	f_1	$F_1 = f_1$	$h_1 = \frac{f_1}{N}$	$H_1 = h_1$
X_2	f_2	$F_2 = f_1 + f_2$	$h_2 = \frac{f_2}{N}$	$H_2 = h_1 + h_2$
X_3	f_3	$F_3 = f_1 + f_2 + f_3$	$h_3 = \frac{f_3}{N}$	$H_3 = h_1 + h_2 + h_3$
...	...	...	...	...
X_i	f_i	$F_i = \sum_{j=1}^i f_j$	$h_i = \frac{f_i}{N}$	$H_i = \sum_{j=1}^i h_j$
...	...	...	...	...
X_k	f_k	$F_k = \sum_{j=1}^k f_j = N$	$h_k = \frac{f_k}{N}$	$H_k = \sum_{j=1}^k h_j = 1$
Total	N		1	

Ejemplo 11.3 Tabla de frecuencias.

Se agrega a la tabla del ejemplo 11.1 las columnas de frecuencias absoluta acumulada (F_i), relativa (h_i) y relativa acumulada (H_i).

Nº. de llamadas	Frec. abs. (f_i)	Frec. abs. acumulada (F_i)	Frec. rel. (h_i)	Frec. rel. acumulada (H_i)
1	8	8	$\frac{8}{30} = \frac{4}{15} = 0.27$	0.27
2	6	14	$\frac{6}{30} = \frac{1}{5} = 0.2$	0.47
3	7	21	$\frac{7}{30} = 0.23$	0.70
4	4	25	$\frac{4}{30} = \frac{2}{15} = 0.13$	0.83
5	5	30	$\frac{5}{30} = \frac{1}{6} = 0.17$	1.00
Total	30		1.00	

La interpretación de la cuarta fila de esta tabla sería: que respecto al número de llamadas diarias en el mes, en 4 días en particular del mes que se está analizando se hicieron 4 llamadas diarias; durante 25 días se hicieron menos de 5 llamadas diarias; en un 13.3% de los días se realizaron 4 llamadas diarias y durante un 83.3% de los 30 días en total se realizaron menos de 5 llamadas diarias.

Tablas de tipo III (Variable cuantitativa continua)

Intervalos de la Variable	Marca de clase X_{MC}	Frec. abs. (f_i)	Frec. abs. acumulada (F_i)	Frec. rel. (h_i)	Frec. rel. acumulada (H_i)
$[a_0, a_1)$	$X_1 = \frac{(a_0 + a_1)}{2}$	f_1	$F_1 = f_1$	$h_1 = \frac{f_1}{N}$	$H_1 = h_1$
$[a_1, a_2)$	$X_2 = \frac{(a_1 + a_2)}{2}$	f_2	$F_2 = f_1 + f_2$	$h_2 = \frac{f_2}{N}$	$H_2 = h_1 + h_2$
$[a_2, a_3)$	$X_3 = \frac{(a_2 + a_3)}{2}$	f_3	$F_3 = f_1 + f_2 + f_3$	$h_3 = \frac{f_3}{N}$	$H_3 = h_1 + h_2 + h_3$
...	...	...	...	...	...
$[a_{i-1}, a_i)$	$X_i = \frac{(a_{i-1} + a_i)}{2}$	f_i	$F_i = \sum_{j=1}^i f_j$	$h_i = \frac{f_i}{N}$	$H_i = \sum_{j=1}^i h_j$
...	...	...	...	...	...
$[a_{k-1}, a_k]$	$X_k = \frac{(a_{k-1} + a_k)}{2}$	f_k	$F_k = \sum_{j=1}^k f_j = N$	$h_k = \frac{f_k}{N}$	$H_k = \sum_{j=1}^k h_j = 1$
Total		N		1	

Estadística y Probabilidades

La marca de clase (X_{MC}) representa los datos que se agrupan en ese intervalo (clase). Es el valor promedio entre los dos límites del intervalo. A diferencia de la tabla anterior, la tabla de intervalo es menos precisa en la determinación de los datos que pertenecen a una clase o intervalo.

Ejemplo 11.4 Tabla de frecuencias.

A la tabla original del ejemplo 11.2 se le agregan tres columnas: frecuencia absoluta acumulada (F_i), frecuencia relativa (h_i) y frecuencia relativa acumulada (H_i).

Intervalos de edades	X_{MC}	f_i	F_i	h_i	H_i
[5, 9)	7	2	2	0.067	0.067
[9, 13)	11	2	4	0.067	0.134
[13, 17)	15	3	7	0.100	0.234
[17, 21)	19	3	10	0.100	0.334
[21, 25)	23	2	12	0.067	0.401
[25, 29)	27	6	18	0.200	0.601
[29, 33)	31	2	20	0.067	0.668
[33, 37)	35	1	21	0.033	0.701
[37, 41)	39	2	23	0.067	0.768
[41, 45)	43	3	26	0.100	0.868
[45, 49)	47	1	27	0.033	0.901
[49, 53)	51	2	29	0.067	0.967
[53, 57]	55	1	30	0.033	1.000
Total		30		1.00	

La interpretación de la cuarta fila de esta tabla sería:

$F_4 = 10$, significa que existen 10 personas con edades comprendidas entre 5 y 21 (de edades mayores o iguales que 5 años y menores que 21 años).

$h_4 = 0.100$, significa que las tres personas cuyas edades están comprendida en el intervalo [17, 21) representan el 10% del total.

$H_4 = 0.334$, significa que el 33.4% de las personas tienen edades comprendidas entre 5 y 21 años.

Ejemplo 11.5 Tabla de frecuencias.

En un estudio realizado a 40 personas acerca del nivel de cotinina, se obtuvieron los siguientes resultados:

1	210	35	103	130	253	123	86
0	44	112	222	234	87	167	284
131	277	477	149	164	121	250	1
173	32	289	313	198	266	245	208
265	3	227	491	17	290	48	173

Construya una tabla de frecuencias.

Solución:

El total de datos con el que se va a trabajar es $N = 40$. El rango $R = 491 - 0 = 491$.

Trabajaremos con 5 intervalos. Por lo cual, la amplitud de los intervalos:

$$i = \frac{491}{5} = 98.2 \approx 100$$

Vamos a redondear a 100 la amplitud de los intervalos por ser un valor conveniente.

Intervalos de la Variable	Marca de clase X_{MC}	Frec. abs. (f_i)	Frec. abs. acumulada (F_i)	Frec. rel. (h_i)	Frec. rel. acumulada (H_i)
[0, 100)	50	11	11	0.275	0.275
[100, 200)	150	12	23	0.300	0.575
[200, 300)	250	14	37	0.350	0.925
[300, 400)	350	1	38	0.025	0.950
[400, 500]	450	2	40	0.050	1
		40		1.00	

Ejemplo 11.6 Tabla de frecuencias.

Determine los valores que hacen falta para completar la siguiente tabla de frecuencias:

Estadística y Probabilidades

Intervalos de la Variable	Frec. abs. (f_i)	Frec. abs. acumulada (F_i)	Frec. relativa (h_i)
[0, 10)	60	60	h_1
[10, 20)	f_2	F_2	0.4
[20, 30)	30	170	h_3
[30, 40)	f_4	F_4	0.1
[40, 50]	f_5	200	h_5
	N		

Solución:

Primero determinemos F_2 . Para esto, se conoce que: $F_3 = f_3 + F_2$, luego:
 $F_2 = 170 - 30 = 140$.

Para calcular f_2 , tenemos que $60 + f_2 = F_2$, por lo que $f_2 = 140 - 60 = 80$.

Como $N = 200$, para calcular f_4 , partimos de la definición de h_4 :

$$h_4 = \frac{f_4}{N} \Rightarrow f_4 = h_4 * N = 0.1 * 200 = 20$$

El valor de F_4 lo determinamos sumando $F_3 + f_4$, así:

$$\begin{aligned} F_4 &= F_3 + f_4 \\ &= 170 + 20 \\ F_4 &= 190 \end{aligned}$$

Siguiendo el mismo procedimiento, determinamos f_5 :

$$\begin{aligned} f_5 &= F_5 - F_4 \\ &= 200 - 190 \\ f_5 &= 10 \end{aligned}$$

Al haber calculado todas las frecuencias absolutas, podemos obtener las frecuencias relativas, dividiéndolas para N :

$$h_1 = \frac{60}{200} = 0.3$$

$$h_3 = \frac{30}{200} = 0.15$$

$$h_5 = \frac{10}{200} = 0.05$$

Escribimos entonces la tabla completa:

Intervalos de la Variable	Frec. abs. (f_i)	Frec. abs. acumulada (F_i)	Frec. relativa (h_i)
[0, 10)	60	60	0.30
[10, 20)	80	140	0.40
[20, 30)	30	170	0.15
[30, 40)	20	190	0.10
[40, 50]	10	200	0.05
	200		1.00

11.3 Gráficos de representación

Objetivos

Al finalizar esta sección el lector podrá:

- * Dada un conjunto de datos, representar la información utilizando histogramas de frecuencias, poligonales de frecuencias y diagramas de tallo y hojas.
- * Interpretar información estadística en forma gráfica.

Al leer un diario o alguna revista, es común encontrarnos con algún tipo de estudio o reportaje cuyos resultados se muestren en tablas y/o gráficos. Es importante, saber representar, leer e interpretar la información que se nos proporciona de esta forma, reconociendo las variables consideradas, tanto lo que se representa en los ejes como el significado de los cambios en los valores de las variables.

Los gráficos permiten formarnos una impresión inmediata acerca del comportamiento de las variables estudiadas, destacando sus características más relevantes.

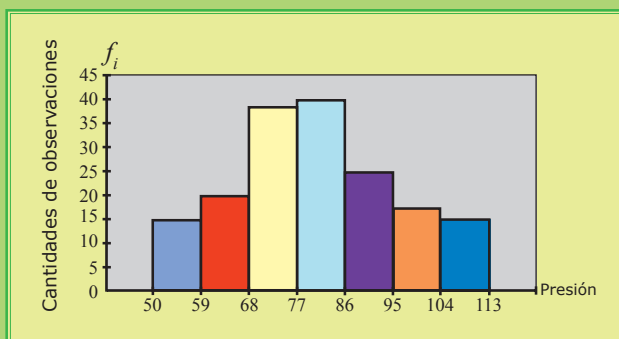
Histograma: Es un gráfico de barras (sin espacios entre ellas), formado por rectángulos, cuya base está dada por la amplitud de cada intervalo y cuyas alturas corresponden a las frecuencias (frecuencias absolutas o relativas) alcanzadas por dichos intervalos.

Estadística y Probabilidades

Ejemplo 11.7 Histograma de frecuencias.

Con los datos organizados en la tabla adjunta, se construye un histograma de frecuencias como se muestra a continuación:

Presión	f_i
[50, 59)	15
[59, 68)	20
[68, 77)	37
[77, 86)	40
[86, 95)	26
[95, 104)	18
[104, 113]	15



Poligonal de frecuencias: Se obtiene al unir las marcas de clases con las frecuencias respectivas, formando un par ordenado. La poligonal formada uniendo estos puntos, se cierra juntando los extremos con las marcas de clase del intervalo anterior al primero y del siguiente al último.

Ejemplo 11.8 Poligonal de frecuencias.

Respecto al ejemplo anterior, la marca de clase de cada intervalo se muestra a continuación:

Presión	f_i	X_{MC}
[50, 59)	15	54.5
[59, 68)	20	63.5
[68, 77)	37	72.5
[77, 86)	40	81.5
[86, 95)	26	90.5
[95, 104)	18	99.5
[104, 113]	15	108.5

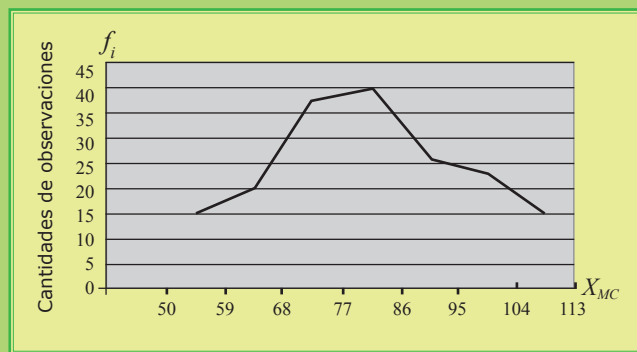


Diagrama de tallo y hojas: Este diagrama permite apreciar la variabilidad o dispersión de los datos.

Ejemplo 11.9 Diagrama de tallo y hojas.

Notas I semestre			Notas II semestre		
2.3	2.0	1.0	5.3	2.0	1.0
4.5	4.6	1.6	4.5	4.6	3.6
2.0	3.0	1.3	6.5	3.0	1.3
5.0	2.7	4.4	5.0	4.7	4.4
1.5	5.5	5.5	4.5	5.5	5.5

Para este caso, la parte entera de la nota constituye el tallo, y las hojas son las cifras decimales de cada nota. En el lado izquierdo están las notas del primer semestre y en el lado derecho están las notas del segundo semestre. En el ejemplo, el gráfico permite la comparación visual de la distribución de las notas por semestre. El diagrama de tallo y hoja es el siguiente:

I Semestre	Tallo	II Semestre
	6	5
0 5 5	5	5 5 3 0
4 5 6	4	7 6 5 5 4
	3	6 0
0 0 3 7	2	0
0 3 5 6	1	3 0
Hojas	Tallo	Hojas

Estadística y Probabilidades

Ejemplo 11.10 Diagrama de tallo y hojas.

En la tabla a continuación se describe el peso en *kg.* de 36 personas.

10.5	8.9	8.7	10.9	15.7	11.1
9.7	15.4	12.3	13.1	15.1	14.5
18.8	12.2	7.6	10.8	17.5	17.6
16.3	13.1	14.3	19.9	14.7	14.9
15.5	18.0	13.5	16.0	18.3	9.4
13.4	17.8	14.6	19.6	15.3	11.5

Construya el diagrama de tallo y hojas.

Solución:

Si seleccionamos como valores del tallo los números 7, 8, 9, ..., 18, 19, el diagrama de tallo y hojas resultante si leemos los datos por filas, sería:

Tallo	Hojas	Frecuencias
7	6	1
8	9, 7	2
9	7, 4	2
10	5, 9, 8	3
11	1, 5	2
12	3, 2	2
13	1, 1, 5, 4	4
14	5, 3, 7, 9, 6	5
15	7, 4, 1, 5, 3	5
16	3, 0	2
17	5, 6, 8	3
18	8, 0, 3	3
19	9, 6	2

11.4 Medidas de tendencia central y no central

Objetivos

Al finalizar esta sección el lector podrá:

- * Dada un conjunto de datos, calcular e interpretar medidas de tendencia central y no central.

Una medida de tendencia central es un número (estadígrafo) que se considera representativo de todos los números en un conjunto de datos.

Media aritmética ($\bar{x}$)

Se define como el cociente entre la suma de los valores que toma la variable (datos) y el total de observaciones: $\bar{x} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n}$; siendo

n el total de observaciones, también se puede expresar como $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$;

generalmente esta definición se ocupa para datos no tabulados. Para datos tabulados en tablas tipo II la media aritmética se obtiene por medio de:

$$\bar{x} = \frac{\sum_{i=1}^k (x_i \cdot f_i)}{n}$$

y en tablas de tipo III se obtiene mediante $\bar{x} = \frac{\sum_{i=1}^k (x_{MC_i} \cdot f_i)}{n}$.

Dentro de las propiedades de la media aritmética se tiene que si se dividen o multiplican todos los datos por una constante, la media aritmética queda dividida o multiplicada, respectivamente, por esa constante; lo mismo que sucede si se le suma o resta una constante.

Sin embargo, la media aritmética presenta el problema de que su valor se puede ver influenciado por valores extremos, que se aparten en exceso del resto de la serie. Estos valores anómalos podrían condicionar en gran medida el valor de la media aritmética, haciéndole perder representatividad.

Ejemplo 11.11 Media aritmética para datos no tabulados.

Se tiene el sueldo de cinco empleados de una empresa: \$567, \$683, \$725, \$675, \$576.

Estadística y Probabilidades

La media aritmética es $\bar{x} = \frac{567 + 683 + 725 + 675 + 576}{5} = \frac{3226}{5} = 645.2$

En este caso, se puede decir que el sueldo promedio que paga la empresa a sus cinco empleados es de \$645.2 .

Ejemplo 11.12 Media aritmética para datos tabulados.

X_{MC} corresponde a la marca de clase del intervalo y se encuentra como la media aritmética de los límites superior e inferior de cada intervalo.

Por ejemplo, la marca de clase para el primer intervalo de la tabla adjunta se encuentra como: $\frac{400 + 450}{2} = 425$

Sueldos	f_i	X_{MC}	$f_i \cdot X_{MC}$
[400, 450)	10	425	4250
[450, 500)	20	475	9500
[500, 550)	30	525	15750
[550, 600)	40	575	23000
[600, 650)	15	625	9375
[650, 700)	10	675	6750
[750, 800]	5	775	3875
	130		72500

La media aritmética es:

$$\bar{x} = \frac{(10)(425) + (20)(475) + (30)(525) + (40)(575) + (15)(625) + (10)(675) + (5)(775)}{130}$$

$$\bar{x} = \frac{4250 + 9500 + 15750 + 23000 + 9375 + 6750 + 3875}{130}$$

$$\bar{x} = \frac{72500}{130} \approx 557.7$$

Mediana ($\tilde{x}$)

Se define como el valor central de una distribución que tiene un número impar de datos, una vez ordenados los datos de manera creciente o decreciente. El dato que representa la mediana divide la distribución en dos grupos, un 50% de valores son inferiores y otro 50% son superiores.

Si el número de datos (N) de la distribución es par, la mediana está dada por el promedio de los dos datos centrales.

La mediana no presenta el problema de estar influenciada por los valores extremos, pero en cambio no utiliza en su cálculo toda la información de la serie de datos (no pondera cada valor por el número de veces que se ha repetido).

Si N es impar, el término central es el dato que ocupa ese lugar:

$$\tilde{x} = x_{\frac{(N+1)}{2}}$$

Ejemplo 11.13 Mediana.

Considere los siguientes datos: 2, 4, 5, 9, 10.

Como N es igual a 5, $\tilde{x} = x_{\frac{(N+1)}{2}} = x_{\frac{(5+1)}{2}} = x_{\frac{6}{2}} = x_3 = 5$.

Si N es par, existen dos datos centrales: $x_{\frac{N}{2}}$ y $x_{\frac{N}{2}+1}$:

$$\tilde{x} = \frac{x_{\frac{N}{2}} + x_{\frac{N}{2}+1}}{2}$$

Ejemplo 11.14 Mediana.

Considere los siguientes datos: 2, 4, 5, 9, 10, 12

Aquí $N = 6$

$$x_{\frac{N}{2}} = x_{\frac{6}{2}} = x_3 = 5; \text{ y,}$$

$$x_{\frac{N}{2}+1} = x_{3+1} = x_4 = 9$$

$$\tilde{x} = \frac{5+9}{2} = 7$$

Moda (Mo)

Se define como el valor de la variable con mayor frecuencia absoluta, o el valor que más se repite. Puede haber más de una moda; para el caso que existan dos, se tiene una distribución bimodal; para más de dos, polimodal.

Estadística y Probabilidades

Ejemplo 11.15 Moda.

Si 60, 75, 75, 80, 90, 90, 100 representan las notas de un estudiante, tenemos 2 modas: 75 y 90. La distribución de estos datos es bimodal.

Ejemplo 11.16 Medidas de tendencia central.

De acuerdo a los datos que se presentan en el ejemplo 11.10, determine las medidas de tendencia central.

Solución:

Como primer paso, ordenemos los 36 datos por columna, lo cual se presenta en la siguiente tabla:

7.6	10.8	13.1	14.6	15.5	17.8
8.7	10.9	13.1	14.7	15.7	18.0
8.9	11.1	13.4	14.9	16.0	18.3
9.4	11.5	13.5	15.1	16.3	18.8
9.7	12.2	14.3	15.3	17.5	19.6
10.5	12.3	14.5	15.4	17.6	19.9

$$\text{Media aritmética } \bar{x} = \frac{7.6 + 8.7 + 8.9 + \dots + 19.6 + 19.9}{36} = 14.069$$

$$\text{La mediana } \tilde{x} = \frac{14.5 + 14.6}{2} = 14.55$$

La moda, el valor de mayor frecuencia, es 13.1 .

Ejemplo 11.17 Media aritmética para datos tabulados.

En la siguiente tabla se describe los resultados agrupados de un estudio a 40 personas sobre el nivel de cotinina:

Nivel de cotinina	f_i	X_{MC}	$f_i \cdot X_{MC}$
[0, 100)	11	50	550
[100, 200)	12	150	1800
[200, 300)	14	250	3500
[300, 400)	1	350	350
[400, 500)	2	450	900
	40		7100

La media aritmética es:

$$x = \frac{(11)(50) + (12)(150) + (14)(250) + (1)(350) + (2)(450)}{40}$$

$$x = \frac{550 + 1800 + 3500 + 350 + 900}{40} = 177.5$$

Ejemplo 11.18 Mediana.

Considere los siguientes datos: 10, 15, 16, x , 20, 22. Si se conoce que la mediana es igual a 17.5, entonces determine el valor del dato x .

Solución:

Como N es igual a 6, una cantidad par, entonces la mediana es:

$$\tilde{x} = \frac{x_{\frac{N}{2}} + x_{\frac{N}{2}+1}}{2} = \frac{16 + x}{2} = 17.5$$

$$16 + x = 35$$

$$x = 19$$

Medidas de tendencia no centrales

También denominadas medidas de localización, permiten conocer otros puntos característicos de la distribución que no son los valores centrales. Entre otros indicadores, se suelen utilizar una serie de valores que dividen la muestra en tramos iguales:

Cuartiles: Son 3 valores que distribuyen la serie de datos, ordenada de forma creciente o decreciente, en cuatro tramos, en los que cada uno de ellos concentra el 25% de los datos.

Deciles: Son 9 valores que distribuyen la serie de datos, ordenada de forma creciente o decreciente, en diez tramos, en los que cada uno de ellos concentra el 10% de los datos.

Percentiles: Son 99 valores que distribuyen la serie de datos, ordenada de forma creciente o decreciente, en cien tramos, en los que cada uno de ellos concentra el 1% de los datos.

Estadística y Probabilidades

Ejemplo 11.19 Medidas de tendencia no centrales.

Vamos a calcular los cuartiles de la serie de datos referidos a la estatura en metros de un grupo de alumnos. Los deciles y percentiles se calculan de igual manera, aunque harían falta distribuciones con mayor número de datos.

Variable (Valor en m)	Frecuencias absolutas		Frecuencias relativas	
	Simple	Acumulada	Simple	Acumulada
1.20	1	1	1/30	1/30
1.21	4	5	4/30	5/30
1.22	4	9	4/30	9/30
1.23	2	11	2/30	11/30
1.24	1	12	1/30	12/30
1.25	2	14	2/30	14/30
1.26	3	17	3/30	17/30
1.27	3	20	3/30	20/30
1.28	4	24	4/30	24/30
1.29	3	27	3/30	27/30
1.30	3	30	3/30	1

1° cuartil: Es el valor $1.22m$, ya que por debajo suyo se sitúa el 25% de la frecuencia acumulada (tal como se puede ver en la columna de la frecuencia relativa acumulada).

2° cuartil: Es el valor $1.26m$, ya que entre este valor y el 1° cuartil se sitúa otro 25% de la frecuencia acumulada.

3° cuartil: Es el valor $1.28m$, ya que entre este valor y el 2° cuartil se sitúa otro 25% de la frecuencia acumulada. Además, por encima suyo queda el restante 25% de esta frecuencia.

Observación: Cuando un cuartil recae en un valor que se ha repetido más de una vez, como ocurre en este ejemplo, la medida de posición no central sería realmente una de las repeticiones.

11.5 Medidas de dispersión

Objetivos

Al finalizar esta sección el lector podrá:

- * Dada un conjunto de datos, calcular e interpretar medidas de dispersión.

Las **medidas de dispersión** estudian la distribución de los valores de la serie, analizando si éstos se encuentran más o menos concentrados, o más o menos dispersos.

Existen diversas medidas de dispersión, entre las más utilizadas podemos destacar las siguientes:

Rango: Mide la amplitud de los valores de la muestra y se calcula por la diferencia entre el mayor valor y el menor.

Varianza: Mide la distancia existente entre los valores de la serie y la media aritmética. Se calcula como la sumatoria de las diferencias al cuadrado entre cada valor y la media, multiplicadas por el número de veces que se ha repetido cada valor. La sumatoria obtenida se divide por el tamaño de la muestra.

$$S^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2 * n_i}{n}$$

Si el número de datos es pequeño, es recomendable utilizar $(n - 1)$ en vez de n en el denominador de la expresión anterior.

La varianza siempre será mayor que cero. Mientras más se aproxima a cero, más concentrados están los valores de la serie alrededor de la media. Por el contrario, mientras mayor sea la varianza, más dispersos están.

Desviación típica o estándar: Se calcula como la raíz cuadrada de la varianza.

Ejemplo 11.20 Medidas de dispersión.

Se tiene una serie de datos de la estatura en metros de los alumnos de una clase y vamos a calcular sus medidas de dispersión.

Estadística y Probabilidades

Variable (Valor en m)	Frecuencias absolutas		Frecuencias relativas	
	Simple	Acumulada	Simple	Acumulada
1.20	1	1	1/30	1/30
1.21	4	5	4/30	5/30
1.22	4	9	4/30	9/30
1.23	2	11	2/30	11/30
1.24	1	12	1/30	12/30
1.25	2	14	2/30	14/30
1.26	3	17	3/30	17/30
1.27	3	20	3/30	20/30
1.28	4	24	4/30	24/30
1.29	3	27	3/30	27/30
1.30	3	30	3/30	1

Rango: Diferencia entre el mayor valor de la muestra (1.30) y el menor valor (1.20). Luego el rango de esta muestra es 0.1m o 10cm.

Varianza: La media aritmética de esta muestra es 1.253m. Luego, aplicamos la expresión:

$$S^2 = \frac{((1.20-1.253)^2*1)+((1.21-1.253)^2*4)+((1.22-1.253)^2*4)+...+((1.30-1.253)^2*3)}{30}$$

Por lo tanto, la varianza es 0.0010 .

Desviación típica: Es la raíz cuadrada de la varianza.

$$S = \sqrt{S^2}$$

Luego:

$$S = \sqrt{0.0010} = 0.0320$$

11.6 Probabilidades

Objetivos

Al finalizar esta sección el lector podrá:

- * Explicar los elementos que definen un espacio muestral.

- * Dadas las condiciones de un experimento aleatorio, describir su espacio muestral.
- * Dadas las condiciones de un experimento aleatorio, calcular la probabilidad clásica de que ocurra un evento requerido.

La rama de la matemática conocida actualmente como probabilidad consiste en el estudio de ciertos experimentos llamados aleatorios, es decir, libres de determinación previa. La probabilidad surgió a comienzos del siglo XVI, en relación con los diversos juegos de azar que se practicaban en la época, más aún, desde civilizaciones tan antiguas como las de los sumerios y egipcios. Hoy en día, los juegos de azar están presentes en aquellos como loterías, juegos de casino y de naipes, muy populares en nuestra sociedad actual. Así, el cálculo de probabilidades determina las posibilidades de ganar o perder en un evento específico. En general, el estudio de las probabilidades permite el análisis de resultados relacionados con fenómenos de carácter social, político y económico, entre otros.

Un **experimento aleatorio** es aquel que está regido por el azar, es decir, se conocen todos los resultados posibles, pero no es posible tener certeza de cuál será en particular el resultado del experimento.

Ejemplo 11.21 Experimentos aleatorios.

- Lanzamiento de un dado.
- Lanzamiento de una moneda.
- Lanzamiento de dos monedas.
- Extracción de una carta de un mazo de naipes.

Se denomina **espacio muestral** (Ω) asociado a un experimento aleatorio, al conjunto de todos los resultados posibles de dicho experimento.

Ejemplo 11.22 Espacios muestrales.

- Al lanzar un dado, el espacio muestral es:

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

- Al lanzar una moneda, el espacio muestral es:

$$\Omega = \{c, s\}$$

Donde: c: cara
 s: sello

Estadística y Probabilidades

- Al lanzar dos monedas, el espacio muestral es:

$$\Omega = \{(c, c), (c, s), (s, c), (s, s)\}$$

Observe que el resultado $(c, s) \neq (s, c)$, es decir, es importante el orden.

- Al extraer una carta de un mazo de naipes, el espacio muestral consta de 52 elementos.

Se denomina **evento** o **suceso** a todo subconjunto de un espacio muestral.

Ejemplo 11.23 Eventos o sucesos.

En el espacio muestral $\Omega = \{1, 2, 3, 4, 5, 6\}$, relacionado con el lanzamiento de un dado, los siguientes son eventos:

- Obtener un número primo:
 $A = \{2, 3, 5\}$
- Obtener un número primo y par:
 $B = \{2\}$
- Obtener un número mayor o igual que 5:
 $C = \{5, 6\}$

Dos eventos son **mutuamente excluyentes** si no pueden ocurrir en forma simultánea, es decir, si y sólo si su intersección es vacía.

Ejemplo 11.24 Eventos mutuamente excluyentes.

- En el lanzamiento de un dado, los eventos:

A: obtener un número par.
B: obtener el número 3.

A y B son mutuamente excluyentes.

El espacio muestral es $\Omega = \{1, 2, 3, 4, 5, 6\}$, y los eventos son:
 $A = \{2, 4, 6\}$ y $B = \{3\}$.

Como $A \cap B = \emptyset$, A y B son mutuamente excluyentes.

- En el lanzamiento de una moneda, los eventos:

A: obtener cara.

B: obtener sello.

A y B son mutuamente excluyentes.

El espacio muestral es $\Omega = \{c, s\}$, y los eventos son $A = \{c\}$ y $B = \{s\}$.

Como $A \cap B = \emptyset$, A y B son mutuamente excluyentes.

En este último caso, además $A \cup B = \Omega$, por lo que A y B se denominarán **eventos complementarios**, es decir, un evento es el complemento del otro:

$$A^c = B \text{ y } B^c = A$$

Probabilidad clásica

Si en un experimento aleatorio todos los resultados son equiprobables (igual probabilidad), es decir, la ocurrencia de uno es igualmente posible que la ocurrencia de cualquiera de los demás, entonces la **probabilidad de un evento** A es la razón:

$$P(A) = \frac{\text{Números de casos favorables para A}}{\text{Números total de casos posibles}}$$

Ésta es la **probabilidad clásica o de Laplace**. A partir de ella, las probabilidades de los posibles resultados del experimento se pueden determinar **a priori**, es decir, sin realizar la experiencia.

Ejemplo 11.25 Cálculo de probabilidades.

Un experimento consiste en lanzar dos dados. Se pide realizar lo siguiente:

- a) Elaborar el espacio muestral del experimento.
- b) Hallar la probabilidad de que al lanzar los dos dados, la suma de las caras de los dados sea igual a 10.
- c) Hallar la probabilidad de que al lanzar los dos dados, la suma se encuentre entre 7 y 9, inclusive.

Solución:

Estadística y Probabilidades

a)

$$\Omega = \left\{ \begin{array}{l} (1, 1), (1, 2), (1, 3), (1, 4), (1, 5), (1, 6), \\ (2, 1), (2, 2), (2, 3), (2, 4), (2, 5), (2, 6), \\ (3, 1), (3, 2), (3, 3), (3, 4), (3, 5), (3, 6), \\ (4, 1), (4, 2), (4, 3), (4, 4), (4, 5), (4, 6), \\ (5, 1), (5, 2), (5, 3), (5, 4), (5, 5), (5, 6), \\ (6, 1), (6, 2), (6, 3), (6, 4), (6, 5), (6, 6), \end{array} \right.$$

b) Sea A: El evento que la suma de las caras de los dados sea igual a 10.

Las combinaciones que cumplen con esta condición son:

$$A = \{(4, 6), (5, 5), (6, 4)\}.$$

$$P(A) = \frac{3}{36} \Leftrightarrow 8.33 \%$$

c) Sea B: El evento que la suma de las caras de los dados se encuentre entre 7 y 9 inclusive.

Las combinaciones que cumplen con esta condición son: $B = \{(1, 6), (2, 5), (2, 6), (3, 4), (3, 5), (3, 6), (4, 3), (4, 4), (4, 5), (5, 2), (5, 3), (5, 4), (6, 1), (6, 2), (6, 3)\}.$

$$P(B) = \frac{15}{36} = \frac{5}{12} \Leftrightarrow 41.67 \%$$

Ejemplo 11.26 Cálculo de probabilidades.

Una comisión ecuatoriana está formada por 20 personas: 8 representantes de la Sierra, 5 de la Costa, 4 del Oriente y 3 de la región Insular. Hallar la probabilidad de seleccionar una persona y que ésta sea:

- a) De la Sierra.
- b) De la Costa.
- c) Del Oriente o de la región Insular.

Solución:

a) Sea A: El evento de seleccionar una persona de la Sierra entre los miembros de la comisión.

$$P(A) = \frac{8}{20} = \frac{2}{5} \Leftrightarrow 40 \%$$

- b) Sea B: El evento de seleccionar una persona de la Costa entre los miembros de la comisión.

$$P(B) = \frac{5}{20} = \frac{1}{4} \Leftrightarrow 25 \%$$

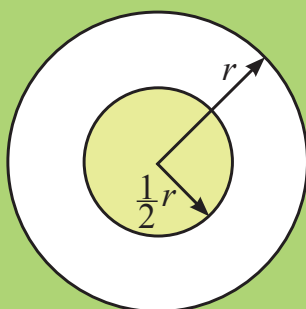
- c) Sea C: El evento de seleccionar una persona del Oriente o de la región Insular.

Como existen 4 personas del Oriente y 3 de la región Insular, por el principio aditivo, existen 7 formas diferentes de seleccionar una persona entre éstas.

$$P(C) = \frac{4+3}{20} = \frac{7}{20} \Leftrightarrow 35 \%$$

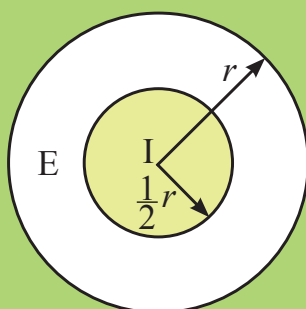
Ejemplo 11.27 Cálculo de probabilidades.

Dada la siguiente figura, hallar la probabilidad de seleccionar un punto al azar que se encuentre en la región sombreada.



Solución:

Si denotamos por E al conjunto de los puntos interiores al círculo de longitud de radio r y denotamos por I al conjunto de los puntos interiores al círculo concéntrico de longitud de radio $\frac{1}{2}r$, tal como se muestra en la siguiente figura:



Estadística y Probabilidades

Sea A: El evento de seleccionar un punto al azar que se encuentre en I.

$$P(A) = \frac{\text{área de I}}{\text{área de E}} = \frac{\pi\left(\frac{1}{2}r\right)^2}{\pi r^2} = \frac{1}{4} \Leftrightarrow 25 \%$$

11.7 Conjuntos y probabilidades

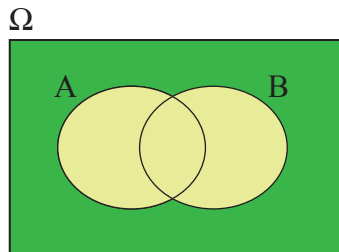
Objetivos

Al finalizar esta sección el lector podrá:

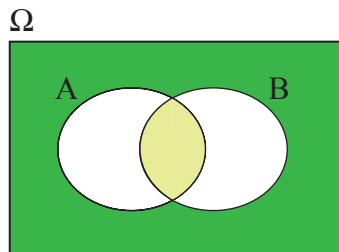
- * Aplicar la teoría de conjuntos al cálculo de probabilidades.
- * Calcular probabilidades de eventos dependientes e independientes.
- * Dadas las condiciones de un experimento aleatorio, calcular la probabilidad de un evento requerido mediante un diagrama de árbol o un triángulo de Pascal.

Conociendo que Ω constituye un espacio muestral, se pueden describir los resultados posibles de un experimento aleatorio en términos de conjuntos, de la siguiente manera:

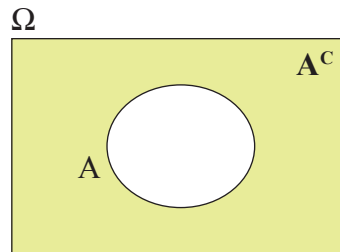
- $A \cup B$: Al menos uno de los eventos A o B ocurre.



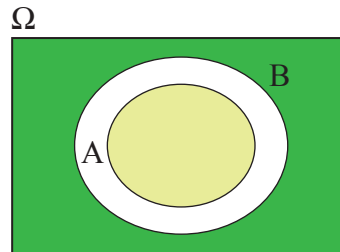
- $A \cap B$: Ambos eventos ocurren.



- A^c : El evento A no ocurre.



- $A \subset B$: Si el evento A ocurre, entonces el evento B también ocurre.



Ejemplo 11.28 Probabilidad.

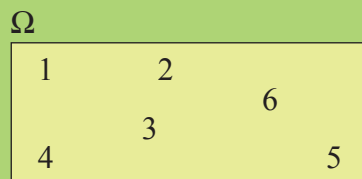
En el experimento del lanzamiento de un dado, sean:

Evento A: Obtener un número par.

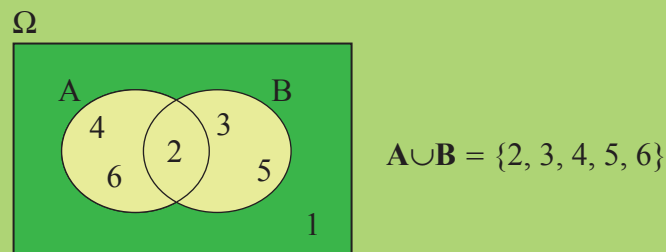
Evento B: Obtener un número primo.

Evento C: Obtener un número distinto de 3 ó 5.

El espacio muestral es $\Omega = \{1, 2, 3, 4, 5, 6\}$.

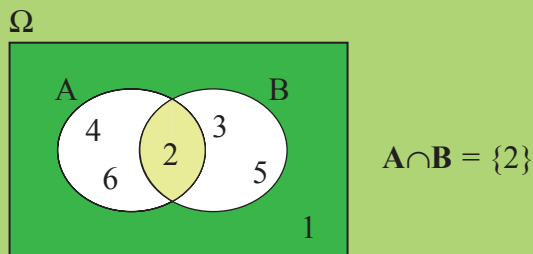


Obtener un número par o un número primo; se puede representar mediante el conjunto $A \cup B$, cuyo diagrama de Venn es:

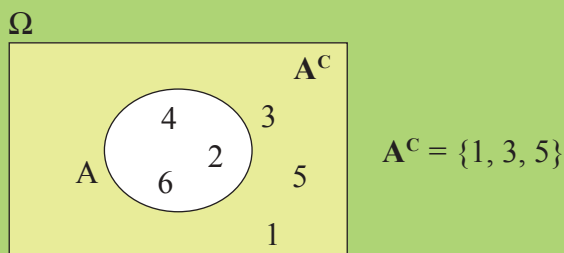


Estadística y Probabilidades

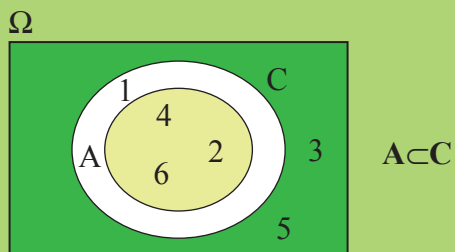
Obtener un número par y primo; se puede representar mediante el conjunto $A \cap B$, cuyo diagrama de Venn es:



No obtener un número par; se puede representar mediante el conjunto A^c , cuyo diagrama de Venn es:



Además, como $A \subset C$, todos los elementos de A son también elementos de C.



Si Ω es un conjunto de n elementos y A un subconjunto de k elementos, entonces la probabilidad de A es:

$$P(A) = \frac{k}{n}$$

Propiedades

- $P(\Omega) = 1$

En el lanzamiento de un dado, para el evento A: obtener un número natural menor que 7, el conjunto es $A = \Omega = \{1, 2, 3, 4, 5, 6\}$ y su probabilidad es:

$$P(A) = \frac{6}{6} = 1$$

- $P(\emptyset) = 0$

En el lanzamiento de un dado, para el evento B: obtener un múltiplo de 4 mayor que 6, el conjunto $B = \{\} = \emptyset$, y su probabilidad es:

$$P(B) = P(\emptyset) = \frac{0}{6} = 0$$

- $0 \leq P(A) \leq 1$

La probabilidad de un evento A es un número real entre 0 y 1.

- $\text{Si } A \cap B = \emptyset$ (A y B se excluyen mutuamente), entonces:

$$P(A \cup B) = P(A) + P(B)$$

En el experimento de extraer una carta al azar de un mazo normal de 52 cartas, los eventos A: Obtener un 5 y B: Obtener un AS, no pueden ocurrir simultáneamente, entonces decimos que son mutuamente excluyentes. La probabilidad de que ocurra A o B, es decir, la probabilidad de que salga 5 o AS es:

$$P(A \cup B) = P(A) + P(B) = \frac{4}{52} + \frac{4}{52} = \frac{8}{52}$$

$$P(A \cup B) = \frac{2}{13}$$

- $P(A^c) = 1 - P(A) \Leftrightarrow P(A) + P(A^c) = 1 \Leftrightarrow P(A) = 1 - P(A^c)$

En el experimento de sacar una carta, el evento no obtener un AS, es el complemento del evento A: obtener un AS, entonces resulta más simple calcular la probabilidad de A^c como $1 - P(A)$:

$$P(A^c) = 1 - P(A)$$

$$= 1 - \frac{4}{52}$$

$$= 1 - \frac{1}{13}$$

$$P(A^c) = \frac{12}{13}$$

Estadística y Probabilidades

- Si $A \cap B \neq \emptyset$ entonces $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

En el experimento del lanzamiento de un dado, los eventos

A: Obtener un número par, y B: Obtener un número primo,

son eventos no disjuntos, ya que $A \cap B = \{2\}$, entonces la probabilidad del evento A o B es:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$= \frac{3}{6} + \frac{2}{6} - \frac{1}{6}$$

$$P(A \cup B) = \frac{4}{6}$$

- Si A y B son **eventos independientes**, es decir, que la ocurrencia de A no influye en la ocurrencia de B, entonces:

$$P(A \cap B) = P(A)P(B)$$

Si lanzamos un dado dos veces, los eventos A: Obtener un número par en el primer lanzamiento, y B: Obtener 1 en el segundo lanzamiento, son eventos independientes, puesto que la probabilidad de un resultado en el segundo lanzamiento no cambia por lo que haya resultado en el primero, entonces, la probabilidad de: Obtener un número par en el primer lanzamiento y 1 en el segundo, es:

$$P(A \cap B) = P(A)P(B)$$

$$= \left(\frac{3}{6}\right) \frac{1}{6}$$

$$P(A \cap B) = \frac{1}{12}$$

- Si A y B son **eventos independientes**, entonces el evento A/B se lee: Ocurre el evento A dado que ha ocurrido B, y tiene la siguiente probabilidad:

$$P(A/B) = P(A)$$

Es decir, la probabilidad de que ocurra el evento A, dado que ha ocurrido B, es igual a la probabilidad de que ocurra el evento A, puesto que A y B son independientes.

Si lanzamos una moneda dos veces consecutivas, la probabilidad $P(A/B)$, sabiendo que los eventos son:

A: Obtener cara en el segundo lanzamiento, y B: Obtener sello en el primero, es:

$$P(A/B) = P(A) = \frac{1}{2}$$

- Si A y B son **eventos dependientes**, es decir, la ocurrencia de A influye en la ocurrencia de B, entonces:

$$P(A \cap B) = P(A) \cdot P(B/A)$$

$$P(B/A) = \frac{P(A \cap B)}{P(A)}$$

Que representa la probabilidad del evento B, dado que ha ocurrido A.

En el experimento de la extracción de una carta: ¿cuál es la probabilidad de que la carta extraída sea la reina de trébol, dado que la carta extraída es también de trébol?

Queremos calcular $P(\text{reina/trébol})$

La probabilidad de "reina y trébol" es $\frac{1}{52}$

La probabilidad de "trébol" es $\frac{13}{52}$

$$\begin{aligned} \text{Así pues, } P(\text{reina/trébol}) &= \frac{P(\text{reina/trébol})}{P(\text{trébol})} \\ &= \frac{\frac{1}{52}}{\frac{13}{52}} \end{aligned}$$

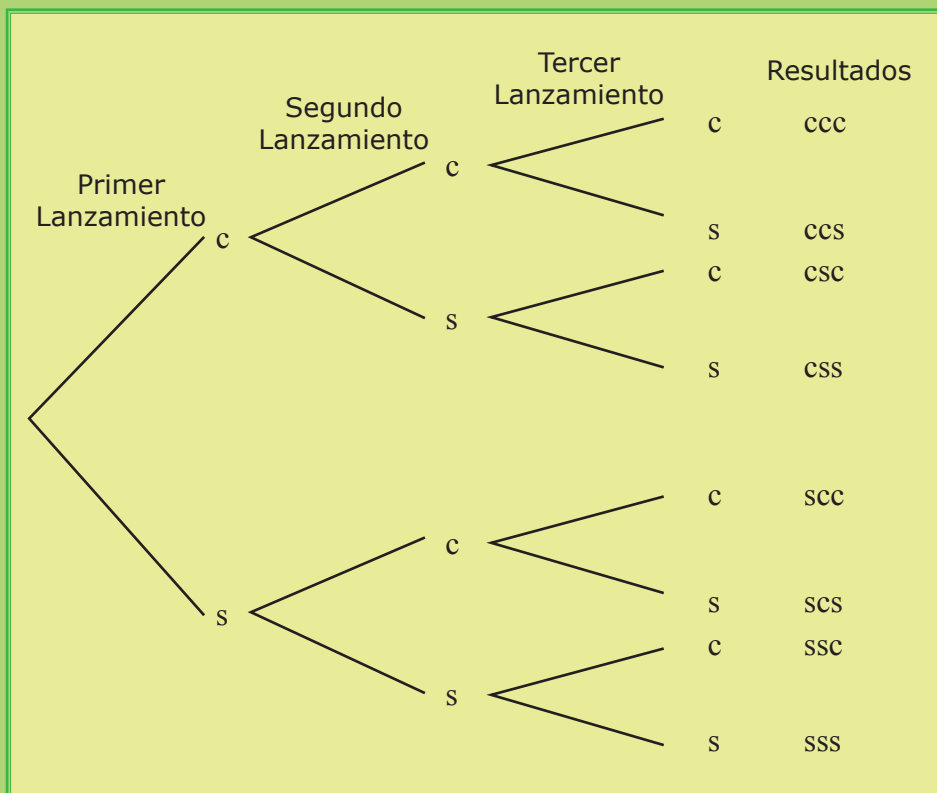
$$P(\text{reina/trébol}) = \frac{1}{13}$$

Para ciertos experimentos aleatorios, se pueden representar de manera gráfica todos los resultados posibles del experimento, que consiste en la iteración de un experimento sencillo, mediante un **diagrama de árbol**. Este tipo de representación permite obtener una técnica de conteo.

Estadística y Probabilidades

Ejemplo 11.29 Diagrama de árbol.

Calcular la probabilidad de obtener dos veces cara y una vez sello, al lanzar tres veces seguidas una moneda.



Los resultados posibles son 8:

ccc – ccs – csc – css – scc – scs – ssc – sss.

Los casos favorables son 3: ccs – csc – scc.

Entonces, la probabilidad buscada es: $\frac{3}{8}$.

Ejemplo 11.30 Cálculo de probabilidades.

Se seleccionan al azar dos cartas entre diez cartas numeradas del 1 a 10. Hallar la probabilidad de que la suma sea impar si:

- a) Se sacan una tras otra sin sustitución.
- b) Las dos cartas se sacan una después de la otra con sustitución.

Solución:

- a) Sea A: El evento en el que la suma sea impar si las dos cartas se sacan una tras otra sin sustitución.

De acuerdo al principio multiplicativo, hay $(10)(9) = 90$ maneras diferentes de sacar dos cartas, una primero que la otra sin sustitución. Como existen $(5)(5) = 25$ maneras de escoger primero un número par y luego uno impar, también existe 25 maneras de escoger primero un número impar y luego uno par; tenemos:

$$P(A) = \frac{25 + 25}{90} = \frac{50}{90} \Leftrightarrow 56\%$$

- b) Sea B: El evento en el que la suma sea impar si las dos cartas se sacan una tras otra con sustitución.

Hay $(10)(10) = 100$ maneras diferentes de seleccionar dos cartas, una después de la otra con sustitución. Según lo mencionado en el literal (b) existen $(5)(5) = 25$ maneras de sacar un número par y luego uno impar, y $(5)(5) = 25$ maneras de sacar un número impar y luego uno par; entonces $P(B) = \frac{25 + 25}{100} = \frac{50}{100} = \frac{1}{2} \Leftrightarrow 50\%$.

Ejemplo 11.31 Cálculo de probabilidades.

Para inventario se asignan códigos numéricos de productos de cuatro cifras a cada artículo existente en una bodega. Suponiendo que todos los números son agotados en la numeración, y si se escogiera un artículo al azar, determine:

- a) La probabilidad de que los dos últimos dígitos asignados a ese artículo sean pares.
- b) La probabilidad de que los dígitos asignados a dicho artículo sean diferentes.
- c) La probabilidad de que los dígitos asignados a dicho artículo sean impares.

Estadística y Probabilidades

Solución:

El número de códigos diferentes para los artículos $(10)(10)(10)(10) = 10000$.

- a) Sea A: El evento en el que los dos últimos dígitos asignados a ese artículo sean pares.

El número posible de códigos con los dos últimos dígitos pares, aplicando el principio multiplicativo, es: $(10)(10)(5)(5) = 2500$.

$$P(A) = \frac{2500}{10000} = 0.25 \Leftrightarrow 25\%$$

- b) Sea B: El evento en el que los dígitos asignados a dicho artículo sean diferentes.

El número posible de códigos con dígitos diferentes, aplicando el principio multiplicativo es $(10)(9)(8)(7) = 5040$.

$$P(A) = \frac{5040}{10000} = 0.504 \Leftrightarrow 50.4\%$$

- c) Sea C: El evento en el que el código de los artículos tenga asignado sólo dígitos impares.

El número posible de códigos con dígitos impares aplicando el principio multiplicativo, es $(5)(5)(5)(5) = 625$.

$$P(C) = \frac{625}{10000} = 0.0625 \Leftrightarrow 6.25\%$$

Ejemplo 11.32 Cálculo de probabilidades.

Se tienen 15 calculadoras, de las cuales 5 son defectuosas. Si se escogen 3 al azar, hallar la probabilidad de que:

- Ninguna sea defectuosa.
- Una exactamente sea defectuosa.
- Una por lo menos sea defectuosa.

Solución:

Existen $C_3^{15} = \binom{15}{3} = 455$ maneras diferentes de escoger 3 calculadoras entre las 15 existentes.

a) Sea A: El evento de no escoger calculadoras defectuosas.

Puesto que hay $15 - 5 = 10$ calculadoras no defectuosas, entonces hay $C_3^{10} = \binom{10}{3} = 120$ maneras diferentes de escoger 3 calculadoras no defectuosas. Así que:

$$P(A) = \frac{120}{455} = \frac{24}{91} \Leftrightarrow 26\%$$

b) Sea B: El evento de escoger exactamente una calculadora defectuosa.

Hay 5 calculadoras defectuosas y $C_2^{10} = \binom{10}{2} = 45$ maneras diferentes de escoger calculadoras no defectuosas; por el principio multiplicativo hay $(5)(45) = 225$ maneras de escoger 3 calculadoras, de las cuales sólo una es defectuosa. Entonces :

$$P(B) = \frac{\binom{5}{1}\binom{10}{2}}{455} = \frac{225}{455} = \frac{45}{91} \Leftrightarrow 49\%$$

c) El evento en que por lo menos una sea defectuosa, es el complemento del evento en que ninguna es defectuosa, entonces:

$$P(A^c) = 1 - P(A) = 1 - \frac{24}{91} = \frac{67}{91} \Leftrightarrow 74\%$$

Ejemplo 11.33 Cálculo de probabilidades.

Seis parejas de casados se encuentran en un cuarto. Si se escogen 2 personas al azar, hallar la probabilidad de que:

- a) Sean esposos
- b) Una sea hombre y otra sea mujer.

Solución:

Existen $C_2^{12} = \binom{12}{2} = 66$ maneras diferentes de escoger 2 personas de las 12.

Estadística y Probabilidades

a) Sea A: El evento de seleccionar una pareja de casados.

Hay 6 parejas de casados, entonces existen $C_1^6 = \binom{6}{1} = 6$ maneras diferentes de escoger a una pareja de casados; por lo tanto:

$$P(A) = \frac{6}{66} = \frac{1}{11} \Leftrightarrow 9\%.$$

b) Sea B: El evento de seleccionar dos personas, tal que una sea hombre y la otra sea mujer.

Como hay 12 personas, 6 hombres y 6 mujeres, hay 6 maneras diferentes de escoger un hombre y 6 maneras diferentes de escoger una mujer; por consiguiente:

$$P(B) = \frac{(6)(6)}{66} = \frac{6}{11} \Leftrightarrow 55\%.$$

Ejemplo 11.34 Cálculo de probabilidades.

En un concurso de belleza conformado por 10 participantes, 3 tienen ojos azules. Si se escogen dos participantes al azar, hallar la probabilidad de que:

- a) Las dos tengan ojos azules.
- b) Una tenga ojos azules.

Solución:

a) Sea A: El evento en el que las participantes seleccionadas tengan ojos azules.

Como existen 10 participantes, existen $C_2^{10} = \binom{10}{2} = 45$ maneras de seleccionar a dos de ellas. Debido a que el evento es que las dos participantes seleccionadas tengan ojos azules, su probabilidad sería:

$$P(A) = \frac{\binom{3}{2}}{\binom{10}{2}} = \frac{3}{45} = \frac{1}{15} \Leftrightarrow 6.67\%.$$

- b) Sea B: El evento en el que una de las participantes tenga ojos azules.

$$P(B) = \frac{\binom{3}{1}\binom{7}{1}}{\binom{10}{2}} = \frac{21}{45} = \frac{7}{15} \Leftrightarrow 46.6\%.$$

Ejemplo 11.35 Cálculo de probabilidades.

Un grupo de amigos, tres hombres y tres mujeres, deciden ir a ver una película; una vez en la sala de cine, consideran sentarse todos en una sola fila. Hallar la probabilidad de que:

- a) Las tres mujeres se sienten juntas.
- b) Los hombres y las mujeres se sienten alternados.

Solución:

- a) Sea A: El evento en el que las mujeres se sienten juntas.

Existen 6 personas en general, lo cual quiere decir que se necesita una fila con 6 asientos disponibles.

Existen 4 posibilidades de tres asientos consecutivos: (1, 2, 3), (2, 3, 4), (3, 4, 5), (4, 5, 6), que serían los asientos ocupados por las mujeres. Como hay 720, es decir $6!$, maneras diferentes de que se sienten los 6 amigos, y existen $3!$ maneras de que se sienten las 3 mujeres en cada posibilidad de tres asientos consecutivos, y a su vez los 3 hombres tienen $3!$ maneras de sentarse en los tres asientos restantes, entonces:

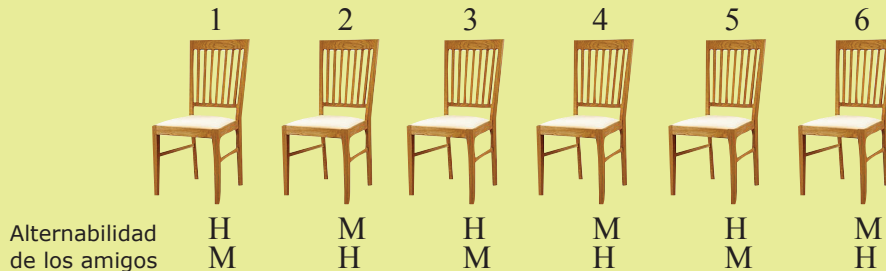
$$P(A) = \frac{(4)(3!)(3!)}{6!} = \frac{1}{5} \Leftrightarrow 20\%.$$

- b) Sea B: El evento en el que las mujeres y los hombres se sienten de forma alternada.

Hay $6! = 720$ maneras de sentarse los seis amigos en los asientos disponibles. De acuerdo al orden alternado y a la ubicación, podemos plantear la situación en la siguiente figura.

Estadística y Probabilidades

Asientos:



Como se puede observar en la figura, existen 2 formas distintas de alternar a los amigos, en las cuales les corresponde 3 asientos a los hombres y 3 a las mujeres; entonces existen 3! maneras diferentes de sentarse, tanto para las mujeres como para los hombres, por lo tanto:

$$P(B) = \frac{(2)(3!)(3!)}{6!} = \frac{72}{720} = \frac{1}{10} \Leftrightarrow 10\%.$$

Ejemplo 11.36 Probabilidad Condicional.

En cierto curso, 25% de los estudiantes reprobaron matemáticas, 15% reprobaron física y 10% reprobaron ambas materias. Se selecciona un estudiante al azar:

- Si reprobó física, ¿cuál es la probabilidad de que reprobó matemáticas?
- Si reprobó matemáticas, ¿cuál es la probabilidad de que reprobó física?
- ¿Cuál es la probabilidad de que reprobó matemáticas o física?

Solución:

Sea $M = \{\text{estudiantes que reprobaron matemáticas}\}$ y
 $F = \{\text{estudiantes que reprobaron física}\}$; entonces:

$$P(M) = 0.25 \quad P(F) = 0.15 \quad P(M \cap F) = 0.10$$

- a) La probabilidad de que el estudiante reprobara matemáticas, dado que haya reprobado física, es:

$$P(M/F) = \frac{P(M \cap F)}{P(F)} = \frac{0.10}{0.15} = \frac{2}{3} \Leftrightarrow 67\%.$$

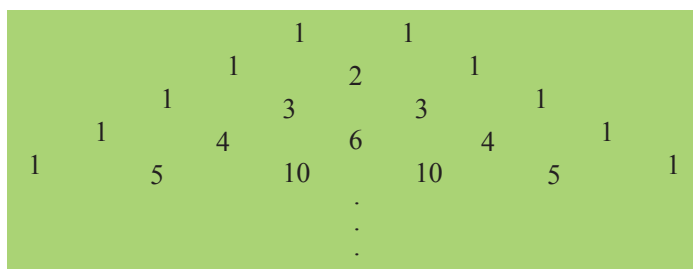
- b) La probabilidad de que el estudiante reprobara física, dado que haya reprobado matemáticas, es:

$$P(F/M) = \frac{P(F \cap M)}{P(M)} = \frac{0.10}{0.25} = \frac{2}{5} \Leftrightarrow 40\%.$$

c) $P(M \cup F) = P(M) + P(F) - P(M \cap F) = 0.25 + 0.15 - 0.10 = 0.30 = \frac{3}{10} \Leftrightarrow 30\%.$

Triángulo de Pascal

Es un triángulo que representa una regularidad numérica de la siguiente forma:



¿Se da cuenta de cómo se forman los números en cada fila?

Es posible relacionar los números que se obtienen en cada línea del triángulo de Pascal con el número de veces de la ocurrencia de un evento determinado, en un experimento aleatorio sencillo que consiste en dos sucesos equiprobables, tales como el lanzamiento de una moneda (cara/sello), género de un hijo (masculino/femenino), entre otros.

Ejemplo 11.37 Triángulo de Pascal.

Calcular el número de casos en que se obtiene tres caras y un sello al lanzar cuatro veces una moneda.

Solución:

Se puede calcular mediante el triángulo de Pascal. Observe la siguiente sucesión de potencias del binomio $(c + s)$:

Estadística y Probabilidades

$$\begin{array}{lcl}
 (c + s)^1 = & & c + s \\
 (c + s)^2 = & & c^2 + 2cs + s^2 \\
 (c + s)^3 = & & c^3 + 3c^2s + 3cs^2 + s^3 \\
 (c + s)^4 = & & c^4 + 4c^3s + 6c^2s^2 + 4cs^3 + s^4 \\
 \cdot & & \cdot \\
 \cdot & & \cdot \\
 \cdot & & \cdot
 \end{array}$$

Los coeficientes de los términos en el desarrollo de cada una de estas potencias, son los números del triángulo de Pascal. ¿Podría usted desarrollar la siguiente potencia y generalizar sus conclusiones para $(c + s)^n$?

En el desarrollo de $(c + s)^4$, el término $4c^3s$ representa 4 casos favorables para el resultado de tres veces cara (c^3) y una vez sello (s): $cccs - ccsc - csc c - sc c c$.

Análogamente, $6c^2s^2$ representa 6 casos favorables para el resultado de dos veces cara (c^2) y dos veces sello (s^2).

Ejemplo 11.38 Cálculo de probabilidades.

Considere el siguiente experimento aleatorio: lanzar un dado tres veces. ¿Es conveniente apostar a favor o en contra de obtener al menos una vez el 5?

Solución:

Como el experimento consta de tres eventos que son independientes entre sí, el número de casos posibles es $(6)(6)(6) = 216$.

Sea A: ninguna vez se obtiene el 5, el número de casos favorables para este evento es $(5)(5)(5) = 125$

$$P(A) = \frac{125}{216} \approx 0.58 \Leftrightarrow 58\%.$$

Para el evento A^c : Alguna vez se obtiene el 5, su probabilidad es:

$$P(A^c) = 1 - P(A) = 1 - 0.58 = 0.42 \Leftrightarrow 42\%.$$

Por lo tanto, no conviene apostar a favor debido a que la probabilidad de ganar es menor a la de perder.

Ejemplo 11.39 Cálculo de probabilidades.

Consideremos una caja de dos bolitas blancas y tres bolitas negras. Extraemos una de ellas al azar y después otra. Calcule la probabilidad de obtener una bolita blanca y una bolita negra, si:

- a) Después de sacar la primera bolita, ésta se devuelve a la caja.
- b) Después de sacar la primera bolita, ésta no se devuelve a la caja.

Solución:

- a) En este caso, los eventos son independientes.

Número de casos posibles = $(5)(5) = 25$.

Número de casos favorables = $(2)(3) + (3)(2) = 12$.

(una blanca y una negra o una negra y una blanca).

$$P(A) = \frac{12}{25} \Leftrightarrow 48\%.$$

$$\text{O bien, } P(A) = \left(\frac{2}{5}\right)\left(\frac{3}{5}\right) + \left(\frac{3}{5}\right)\left(\frac{2}{5}\right) = \frac{12}{25}.$$

- b) En este caso, los eventos son dependientes.

Número de casos posibles = $(5)(4) = 20$.

Número de casos favorables = $(2)(3) + (3)(2) = 12$.

$$P(B) = \frac{12}{20} = \frac{3}{5} \Leftrightarrow 60\%.$$

$$\text{O bien, } P(B) = \left(\frac{2}{5}\right)\left(\frac{3}{4}\right) + \left(\frac{3}{5}\right)\left(\frac{2}{4}\right) = \frac{3}{5}.$$

Ejemplo 11.40 Cálculo de probabilidades.

En el juego de la ruleta, una bola puede caer en cualquiera de las 37 casillas, que son sectores circulares iguales, numerados del 0 al 36. Hay 18 casillas rojas, 18 negras y una blanca (la del 0). En el juego gana el número de la casilla en que cae la bola, después de hacerla girar libremente y detenerse por sí sola. ¿Cuál es la probabilidad de que la bola caiga en una casilla numerada con un número par mayor que 29?

Estadística y Probabilidades

Solución:

Como la bola puede caer en cualquiera de las casillas, los casos posibles son 37.

Los números (o casillas) mayores que 29 son:

30 – 31 – 32 – 33 – 34 – 35 – 36.

De ellos, son números pares: 30, 32, 34 y 36.

Por lo tanto, los casos favorables son 4 y la probabilidad requerida es 11%.

Ejemplo 11.41 Cálculo de probabilidades.

La probabilidad de que un alumno haya aprobado la parte teórica del examen para obtener licencia para conducir es 0.68; la probabilidad de que haya aprobado el examen práctico es 0.72; y, la probabilidad de que haya aprobado alguna de las dos partes es 0.82. Elegido un alumno al azar, encuentre la probabilidad de que haya aprobado el examen para obtener licencia.

Solución:

Para obtener la licencia, cualquier persona debe aprobar ambas partes del examen: la teórica y la práctica. Nos preguntan entonces, por la probabilidad de que una persona elegida al azar haya aprobado las dos partes del examen.

Si denominamos:

A: Aprobar la parte teórica.

B: Aprobar la parte práctica.

Debemos calcular la probabilidad de A y B, o bien: $P(A \cap B)$. Sabemos que $P(A \cup B) = P(A) + P(B) - P(A \cap B)$. Si despejamos $P(A \cap B)$, tenemos:

$$P(A \cap B) = P(A) + P(B) - P(A \cup B)$$

Del enunciado, tenemos que:

$$P(A) = 0.68, P(B) = 0.72 \text{ y } P(A \cup B) = 0.82.$$

Reemplazando en la ecuación anterior, se tiene:

$$P(A \cap B) = 0.68 + 0.72 - 0.82, \text{ de donde}$$

$$P(A \cap B) = 0.58$$

Por lo tanto, la probabilidad que un alumno elegido al azar obtenga su licencia es 58%.

Ejemplo 11.42 Diagrama de árbol con cálculo de probabilidades.

Existen tres cajas que contienen unidades de CD-ROM, de acuerdo al siguiente detalle:

Caja I: contiene 10 CD-ROM, de los cuales 4 son defectuosos.

Caja II: contiene 6 CD-ROM, de los cuales 1 es defectuoso.

Caja III: contiene 8 CD-ROM, de los cuales 3 son defectuosos.

Si se escoge al azar una caja y luego sacamos al azar un CD-ROM. ¿Cuál es la probabilidad de que el CD-ROM sea defectuoso?

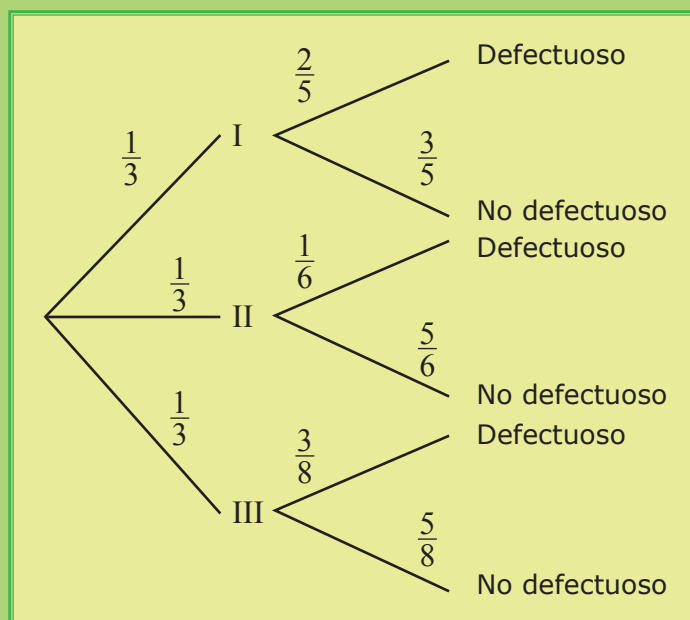
Solución:

Aquí realizamos una serie de dos experimentos consecutivos:

- a) Escoger una de las tres cajas.
- b) Escoger un CD-ROM que sea defectuoso o no defectuoso.

El siguiente diagrama de árbol describe cada experimento consecutivo, en donde cada rama es un experimento y la probabilidad de cada rama del árbol se coloca según cada caso:

Estadística y Probabilidades



La probabilidad de que una trayectoria del árbol suceda es, según el principio multiplicativo, el producto de las probabilidades de cada rama de la trayectoria, o sea, que la probabilidad de escoger la caja I y luego un CD-ROM defectuoso es $\left(\frac{1}{3}\right)\left(\frac{2}{5}\right) = \frac{1}{15}$.

Ahora, como hay tres trayectorias mutuamente exclusivas que conducen a un CD-ROM defectuoso, por el principio aditivo, la suma de las probabilidades de estas trayectorias es la probabilidad buscada, así:

$$p = \left(\frac{1}{3}\right)\left(\frac{2}{5}\right) + \left(\frac{1}{3}\right)\left(\frac{1}{6}\right) + \left(\frac{1}{3}\right)\left(\frac{3}{8}\right) = \frac{2}{15} + \frac{1}{18} + \frac{3}{24} = \frac{48 + 20 + 45}{360} = \frac{113}{360} \Leftrightarrow 31\%.$$

Ejemplo 11.43 Diagrama de árbol con cálculo de probabilidades.

Tres empresas de comercialización, A, B y C, han solicitado a los egresados de Auditoría de la ESPOL, realizar una auditoría de sus estados contables. Se sabe que la empresa A tiene 180 estados contables, mientras que la B y la C tienen 250 y 300, respectivamente. Se sabe también que 18 estados contables de la A tiene errores, 20 de los de B tiene errores y 33 de los de C también tiene errores. Suponiendo que los estudiantes auditores hacen el trabajo correctamente, y se

escoge al azar un estado contable: a) ¿Cuál será la probabilidad de que éste tenga errores? b) Si el estado contable escogido no contiene errores, ¿cuál será la probabilidad de que provenga de la empresa B?

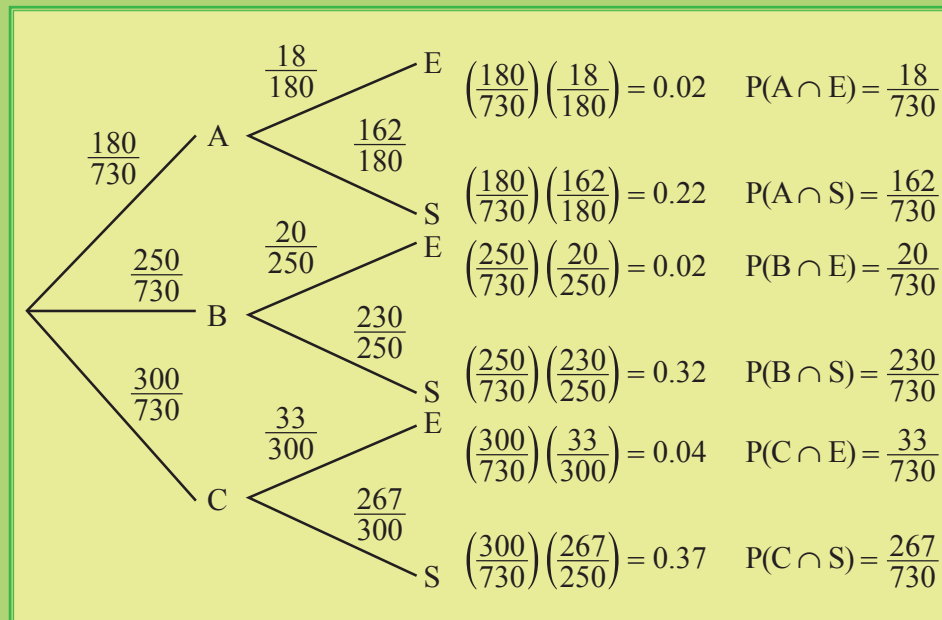
Solución:

Sean los eventos:

A: Seleccionar un balance de la empresa de comercialización A.
 B: Seleccionar un balance de la empresa de comercialización B.
 C: Seleccionar un balance de la empresa de comercialización C.

E: En el que el estado contable seleccionado tenga errores.

S: En el que el estado contable seleccionado no tenga errores.



a) De acuerdo al evento E y por el principio aditivo:

$$P(E) = 0.02 + 0.02 + 0.04 = 0.1 \Leftrightarrow 10\%$$

b) Sea S: El evento en el que el estado contable seleccionado no tenga errores.

$$P(S) = 1 - 0.1 = 0.9 \Leftrightarrow 90\%$$

$$P(B/S) = \frac{P(B \cap S)}{P(S)} = \frac{0.32}{0.9} = 0.35 \Leftrightarrow 35\%.$$

Para los ejercicios de las secciones 11.2, 11.3, 11.4 y 11.5, considere los siguientes conjuntos de datos A, B y C, respectivamente.

A. Edad en años de estudiantes en etapa colegial	B. Notas de los estudiantes de un curso de Cálculo en la ESPOL	C. Tiempo de espera en minutos de los clientes de un banco de la localidad
16	6.45	40
12	8.30	45
12	7.55	52
11	6.00	33
19	8.20	27
16	6.25	5
17	6.00	11
11	7.00	31
16	6.40	42
9	7.45	51
17	6.20	55
16	6.35	55
12	6.55	60
13	7.80	42
10	6.00	37
16	6.45	35
17	7.95	10
14	6.00	43
17	6.15	54
16	7.05	55
18	7.35	10
13	6.25	22
14	6.00	5
13	6.45	62
13	6.60	74
13	9.15	57
14	6.60	42
17	7.60	43
14	6.35	31
12	7.30	26
14	7.40	29
15	8.15	35
14	6.70	33
13	6.25	41
17	6.40	39
10	6.45	44
15	8.30	54
15	7.55	56
12	6.00	22
14	8.20	15
13	6.25	32
10	6.00	17
12	7.00	26
16	6.40	42
12	7.45	44
14	6.20	45

11.2 Organización de datos

1. Con cada uno de los conjuntos de datos dados:
 - a) Construya una tabla de frecuencias estimando un número de intervalos adecuado.
 - b) Complete la tabla con las frecuencias: absoluta, absoluta acumulada, relativa y relativa acumulada.
 - c) Calcule la marca de clase para cada intervalo.
 - d) Formule al menos cuatro interpretaciones expresándolas con sus propias palabras. Discútalas con sus compañeros y analice su validez.

11.3 Gráficos de representación

2. Con cada uno de los conjuntos de datos dados:
 - a) Construya el histograma de frecuencias.
 - b) Construya la poligonal de frecuencias.
 - c) Formule al menos cuatro interpretaciones que puedan ser sustentadas con los gráficos construidos.
3. Con los datos del grupo B, construya un diagrama de tallo y hojas.

11.4 Medidas de tendencia central

4. Para cada uno de los conjuntos de datos dados:
 - a) Calcule los estadísticos de tendencia central.
 - b) Describa los cuartiles y deciles.
 - c) Formule al menos cinco interpretaciones que puedan ser sustentadas con la información obtenida en los literales anteriores.
5. Analice la ubicación de los estadísticos obtenidos en el ejercicio anterior respecto al histograma de frecuencias.

11.5 Medidas de dispersión

6. Para cada uno de los conjuntos de datos dados:
 - a) Calcule los estadísticos de dispersión.
 - b) Formule al menos tres interpretaciones que puedan ser sustentadas con la información obtenida en el literal anterior.

11.6 Probabilidades

7. En el experimento de lanzar dos dados y el resultado es la multiplicación de los dos números obtenidos, determine:
 - a) El espacio muestral.
 - b) La probabilidad de que el resultado sea 12.
 - c) La probabilidad de que el resultado sea menor que 10.
 - d) Discuta si los dos eventos anteriores son equiprobables.

8. En un concurso se deben obtener aleatoriamente dos dígitos del 1 al 5, de manera independiente. Si la suma de los dos dígitos es par, el participante gana \$100; si la suma de los dos dígitos es impar, gana \$10.
- a) Escriba el espacio muestral para este concurso.
 - b) Encuentre la probabilidad de que el participante gane \$100 en un intento.
 - c) ¿Le recomendaría este juego a alguien? ¿Por qué?
9. Se pregunta a tres personas distintas, elegidas al azar, si son partidarias o no de un determinado candidato a la Presidencia de la República.
- a) Determine el espacio muestral asociado a dicho experimento, utilizando la letra "*s*" para las respuestas afirmativas y "*n*" para las negativas.
 - b) Calcule la probabilidad de que al menos una persona es partidaria del candidato.
 - c) Calcule la probabilidad de que ocurra el evento contrario al del literal anterior.
10. Se pregunta a cuatro personas elegidas al azar que consumen carnes, si les gusta la carne de cerdo.
- a) Determine el espacio muestral asociado a dicho experimento, utilizando la letra "*s*" para las respuestas afirmativas y "*n*" para las negativas.
 - b) Calcule la probabilidad de que al menos tres personas gusten de la carne de cerdo.
 - c) Calcule la probabilidad de que, a lo mucho, sólo a dos personas les guste la carne de cerdo.
11. Discuta el siguiente problema: Tengo dos urnas, dos bolas blancas y dos bolas negras. Se desea saber cómo debo distribuir las bolas en las urnas para que, al elegir una urna al azar, sea máxima la probabilidad de obtener una bola blanca. La única condición exigida es que cada urna tenga al menos una bola.

11.7 Conjuntos y Probabilidades

12. Las estadísticas de accidentes de tránsito indican que el 80% de la población ha sufrido algún tipo de accidente. El 60% ha sufrido daños en su vehículo y el 35% ha sufrido daños físicos. Determine la probabilidad de que un conductor sufra los dos tipos de daños.

13. Mediante un estudio de opinión, se estableció que las tres características preferidas por el electorado sobre un candidato a la Alcaldía son:

- Hombre (H)
- Sin afiliación política (N)
- Casado (C)

Se sabe que el 75% de los encuestados prefieren un candidato de género masculino, el 60% lo prefiere sin afiliación política, el 70% prefiere que sea casado, el 90% prefieren que sea hombre o sin afiliación política, el 85% prefiere que sea hombre o casado, el 95% prefiere que sea sin afiliación política o casado, el 99% prefiere alguna de las tres características para el candidato.

Construya un diagrama de Venn para representar las tres características.

Cuál es la probabilidad de que una persona:

- a) Prefiera las tres características.
- b) No prefiera alguna de las tres características.
- c) Prefiera que sólo sea casado.

14. A través de una encuesta se determinó que el 80% de los habitantes se informa diariamente de las noticias viendo la televisión, el 40% se informa por el periódico y el 30% por ambos medios. Determine:

- a) La probabilidad de que una persona de esa ciudad no se informa por alguno de estos medios.
- b) La probabilidad de que una persona que se informó por la televisión haya leído el periódico.
- c) Sabiendo que una persona leyó el periódico, ¿cuál es la probabilidad de que no vio los noticieros?

15. Mediante una encuesta, se ha determinado que el 80% de las compras diarias en cierto supermercado son realizadas por mujeres. De este grupo, el 70% compra algún producto en promoción, mientras que del grupo de los hombres, sólo el 10% compra productos en promoción.

- a) Determine la probabilidad de que el almacén venda diariamente algún producto en promoción.
- b) Si se selecciona una venta al azar y no contiene productos en promoción, ¿cuál es la probabilidad que la compra haya sido hecha por una mujer? ¿Y por un hombre?

16. Cierta cooperativa de ahorros ha observado que la probabilidad de que un préstamo sea cancelado es del 70% si el prestamista es de género femenino. La proporción de ahorristas hombres respecto a mujeres es de 1. Determine la probabilidad de que un ahorrista pague un préstamo y sea de género femenino.

17. Cierta concurso consiste en obtener aleatoriamente un número divisible para 10 de todos los números enteros comprendidos del 1 al 40.

Si el número obtenido es divisible para 10, el participante recibe el llavero A; en caso contrario, recibe el llavero B. A continuación, se escoge una llave. El llavero A contiene 3 llaves de un vehículo y 7 llaves que no corresponden al vehículo. El llavero B tiene 2 llaves del vehículo y 10 que no lo son.

Encuentre:

- a) La probabilidad de que la llave extraída sea del vehículo.
 - b) La probabilidad de que la llave extraída sea del vehículo y pertenezca al llavero B.
18. Cierta empresario requiere decidir sobre la mejor opción para elevar sus ventas. Existen dos situaciones, A y B, que afectan su mercado. La situación A tiene una probabilidad de ocurrir igual a 0.4, y la B de 0.6. En ambas situaciones puede invertir en un anuncio. Si ocurre A, la probabilidad de invertir en un anuncio es de 0.8, y si ocurre B, la probabilidad es de 0.3. El empresario desea invertir en el anuncio. Determine:
- a) La probabilidad de que el empresario invierta en el anuncio.
 - b) La probabilidad de que el empresario no invierta en el anuncio.
 - c) ¿Qué recomendaría usted si el evento A es una situación que favorece sus ventas?
19. Considere que 0.076 es la probabilidad de que cuatro clientes diferentes compren un televisor con pantalla plana en un almacén. Determine:
- a) La probabilidad de que uno de ellos compre un televisor con pantalla plana.
 - b) La probabilidad de que al menos dos de ellos compren un televisor con pantalla plana.
20. Un examen consta de 6 preguntas. Cada pregunta tiene cinco opciones, de las cuales sólo una es la respuesta correcta. El número mínimo de preguntas que un estudiante debe contestar correctamente para aprobar el examen es 4. Determine la probabilidad de que un estudiante apruebe el examen contestando al azar las preguntas.